U. S. NAVAL SUBMARINE MEDICAL CENTER

Submarine Base, Groton, Conn.

REPORT NUMBER 593

AN EFFICIENT MONAURAL PROCEDURE FOR THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

by

J. Donald Harris, Ph.D.

Bureau of Medicine and Surgery, Navy Department

Research Work Unit MF12.524.004-9010D.05

Approved and Released by:

J. E. Stark, CAPT MC USN
COMMANDING OFFICER
Naval Submarine Medical Center

5 September 1969

AN EFFICIENT MONAURAL PROCEDURE FOR THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

by

J. Donald Harris, Ph.D.

SUBMARINE MEDICAL RESEARCH LABORATORY
NAVAL SUBMARINE MEDICAL CENTER REPORT NO. 593

Bureau of Medicine and Surgery, Navy Department
Research Work Unit MF12.524.004-9010D.05

Reviewed and Approved by:

Charles F. Gell

Charles F. Gell, M.D., D.Sc. (Med.)
Scientific Director
SubMedResLab

Reviewed and Approved by:

Joseph D. Bloom.

Joseph D. Bloom, CDR MC USN
Director
SubMedResLab

Approved and Released by:

J. E. Stark

J. E. Stark
CAPT MC U. S. Navy
COMMANDING OFFICER, SMC

SUMMARY PAGE

THE PROBLEM

To explore new and more efficient ways to calibrate the frequency response of an earphone when coupled to the human head.

FINDINGS

A procedure was invented which reduces the listener's task to simple monaural loudness discrimination, eliminates half of the sources of variance inherent in the traditional procedure, and can be accomplished in a fraction of the usual time.

APPLICATION

For electrical engineers, sonar technicians, communications engineers, otologists, audiologists, and others interested in the specification of the real-ear response of an earphone.

ADMINISTRATIVE INFORMATION

This investigation was conducted as a part of Bureau of Medicine and Surgery Research Work Unit MF12.524.004-9010D — "Optimization of Auditory Performance in Submarines." The manuscript for the present report was approved for publication on 5 September 1969, and designated as SMRL Report No. 593. The present report is No. 5 on this Work Unit.

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL CENTER

ABSTRACT

A new procedure is examined for psychoacoustic calibration of earphones in which the air-conducted outputs of a standard and an unknown earphone are successively equated for loudness to a reference bone-conducted tone. The problem to the subject is one of monaural loudness discrimination, with a relatively small variance (differential sensitivity = 1.23 — 1.61 dB), and involves only four sources of variance associated with coupling two earphones to the same ear, and a single loudness discrimination judgment for each phone. The mean test-retest difference in the earphone transfer functions varied by 1.33 — 5.89 dB at different frequencies, mid-value = 3.36 dB. Only a few minutes are required to complete a subject's observations at any frequency. Acceptable group means for transferring audiometric standards to an unknown earphone could be obtained at any frequency by requiring as few as nine subjects to make a single monaural loudness discrimination per earphone by this technique. In contrast, the traditional alternate interaural loudness balancing has a somewhat larger variance associated with the judgment **per se** (differential sensitivity = 1.50 — 2.50 dB), and involves the additional variances associated with collecting two absolute thresholds and coupling the two earphones to a second ear when the earphones are reversed on the head. The mean test-retest difference in the earphone transfer functions by the traditional "ear-reversal" method varied from 4.16 — 7.54 dB at different frequencies, mid-value = 6.30 dB, nearly twice that of the suggested procedure.

AN EFFICIENT MONAURAL PROCEDURE FOR THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

In transferring audiometric threshold SPLs from a standard earphone to an earphone of different sensitivity and physical configuration, it is sometimes unsatisfactory to measure levels generated by the two earphones successively in a closed acoustic coupler, or even in the actual cavity enclosed by an earphone placed on a human head. Where this is the case, as with an insert earphone, or one of the large circumaural earphone/cushion units, a psychoacoustic loudness balance must be performed with a panel of listeners making equal-loudness judgments at each audiometric frequency between a standard and a new earphone. For example, the ISO specification (anon, 1964) for reference equivalent threshold SPLs, given for five different earphones from five different countries, is based partly on psychoacoustic judgments of loudness equality performed by one or more laboratories in each country on the standard earphone in that country compared with those from two other countries (see Weissler, 1968).

Differences amounting to several decibels are seen between the mean transfer functions from different laboratories, relating the voltage in one phone to the voltage on another type of phone which yields equal loudness. Weissler (1968) recounts the final results of a number of loudness balances among the audiometric earphones of five countries; the final estimate of the standard errors of the transfer functions between any two earphones, from at least two countries, was of the order of 4 dB, from which we may conclude that the uncertainty in a substantial number of subjects was considerably larger. The only recouse one has is to mass observers and observations until the standard error of the mean is acceptably small; but this approach is costly in time, and short-cuts are often adopted, to the degradation of the data.

Although no national standards have ever been promulgated of procedures for loudness-balancing between earphones, a convention has been informally followed of alternate interaural loudness balancing of an unknown phone applied to one ear against a standard phone set successively on the other ear at sensation levels of 0, 20, and 40 dB. After judgments of equality between the two ears have been made at the frequencies desired, the subject replaces the phones on the opposite ears, to allow for differences in equal-loudness contours (including threshold, or 0-loundess contour), and renders another series of judgments. Ear differences are then scrubbed out by simple arithmetic, and the voltage noted to the unknown phone which yields equal loudness to a standard voltage in the standard phone.

In performing such loudness balances using some of the newer circumaural phone/cushion units, we became greatly concerned with the variances in the data, with a view ultimately of reducing the sources of such variance so far as possible. Even a cursory glance at the problem reveals as a bare minimum those variances associated with:

(1) (2) coupling of the standard phone to the L, and again to the R ear when the phones are reversed to account for ear acuity differences,

(3) (4) same for the unknown phone on the opposite ear,

(5) (6) the constant, variable, and accidental errors inherent in the taking of absolute threshold with the standard phone on the two ears, and

(7) (8) the errors associated with the two sets of loudness judgments demanded.

The more exact statement of the sources of these and other variances which combine in a total earphone transfer function has never been made, and little quantitative data are at hand on the extent of the variance ascribable to each source. This letter asseses the variance associated strictly with (7) (8), and compares the total variances (1)-(8) of the earphone transfer function by the traditional alternate interaural method with a method using only one ear in a loudness discrimination task in which the variance is less than in (7) (8), and which avoids altogether

sources (3)-(6), for a total of only four instead of eight sources of variance.

EXPERIMENT I: Traditional Earphone Loudness Balancing

A. The Differential Sensitivity for Alternate Interaural Loudness Discrimination.

Twelve graduate students in sensory psychophysiology served as subjects, and an older psycho-acoustician (not the author) with a mild high-tone hearing defect. They adjusted with a 1-dB/step attenuator the loudness of one of two circumaural earphones to equality with a standard W. E. 705A earphone set at 40 dB sensation level. Channel interrupter switches (normally "off") were used by the subjects at will, but never together so as to create a simultaneous binaural condition. Subject had no visual cue for position of the 1-dB/step rotary potentiometer governing the channel to the circumaural earphone; earphones, ear order, and frequency were suitably counterbalanced to drop out audiometric differences in the two ears, fatigue, order effects, etc. Ten consecutive equality judgments were demanded before the earphones were again touched. The standard deviation (SD) of any such series of N judgments is the traditional index to differential sensitivity. Tables I-II give these data for the nine frequencies, for each of two circumaural units. There is no systematic difference between the two ears, and in the last row are given the mid-values for all 26 ears (it is not strictly correct to average these SDs). The mid-values increase progressively from about 1.5 at 250 to about 2.5 dB at 8000 Hz.

No data of just this type have come to our attention, but there exist several sets of data from simultaneous interaural (diotic) loudness judgments (for reviews see Harris (1963) and Rowland and Tobias (1967)). The latter have provided mean sensitivities of 1.15, 0.72, and 0.92 dB at .25, 2, and 6 kHz respectively, as compared with values of 0.88, 0.65, and 0.93 for the monotic condition, at overall loudness comparable to ours. If there is an effect of frequency, it is negligible.

Unfortunately, no estimates of variance were included, so that the precision of such values cannot be estimated and compared with ours. Furthermore, their subjects tracked the presence of intensity modulation in a modification of the Békésy Method of Limits, so that the mean sensitivity for each subject was an average of judgments "just noticeable difference" and "just not noticeable difference"; these are more traditionally termed "jnd" and "jnnd", and their average the JND. It cannot be compared directly with the Differential Threshold (DL) from the Method of Constants, nor to the SD from the Method of Adjustments without appropriate transfer studies, which have never been done completely for loudness discrimination. Thus, the means in Tables I-II are a function both of the underlying sensitivity and of the variant of psychophysical method, but the contribution of each to the mean sensitivity and to its variance cannot be assessed at this time.

It might be supposed that the simultaneous interaural judgment would be more sensitive, and less variable, since an additional cue is present, namely, directionality of the phantom image in phenomenological space. A relatively slight interaural difference in intensity might move the sound image left or right and act as a vernier on the coaser scale. A direct comparison of this possibility was made by Jerger and Hartford (1960) who found no difference in sensitivities by the Method of Adjustment for normal-hearing persons, though with persons with unilateral hypacusis the picture was entirely different. In such persons, differences between the two methods of as much as 10 dB were common. Evidently, the two types of interaural judgments are not at all alike, though they yield the same mean SDs on normal subjects. Jerger and Harford did not furnish any estimate of individual or group variance.

B. Distributions of Earphone Transfer Functions.

Table III gives the SDs of the distributions of the individual transfer functions for the two circumaural earphones, together with estimates of the precision of the mean transfer function. In this traditional balancing procedure, two of three subjects agree with

TABLE I

DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS DISCRIMINATION

Entry: Standard Deviation in DB of Ten Consecutive Loudness Equality Judgments at 40 DB Sensation Level on Standard Earphone. Comparison Earphone: Maico Co. "Auraldome"

Subj.	Stand. Phone On	0.25	0.50	0.75	1	2	3	4	6	8
					Frequency in KHz					
AR	L	3.26	1.89	1.96	1.96	3.26	2.04	2.97	3.29	3.26
	R	.83	2.00	1.73	2.10	2.87	2.11	3.61	2.32	2.28
JR	L	1.22	2.02	1.04	1.68	1.48	2.65	3.26	1.55	3.22
	R	.77	1.60	1.17	1.94	1.00	1.62	1.64	1.68	2.83
JD	L	1.99	.92	2.04	1.04	1.80	1.22	1.28	1.20	.98
	R	1.27	1.04	1.17	1.87	1.95	2.36	1.91	1.37	2.19
JW	L	1.34	2.62	1.27	3.68	2.47	2.30	2.19	1.90	3.35
	R	1.55	1.11	1.37	1.18	1.17	1.87	3.07	3.03	3.23
HM	L	3.69	2.26	1.56	2.00	2.59	2.29	3.54	1.85	2.40
	R	1.11	1.64	2.71	2.74	3.72	1.95	4.85	4.05	2.18
CMc	L	.66	.64	1.37	1.20	1.17	1.58	2.27	2.33	3.46
	R	1.19	.81	1 44	1.25	2.00	2.38	1.47	1.63	1.20
MD	L	2.34	1.47	2.11	2.47	2.19	1.40	3.21	2.24	2.19
	R	1.36	1.80	1.40	2.29	2.29	2.73	5.62	4.45	4.38
DW	L	1.43	2.24	3.01	.92	.80	3.19	2.32	2.90	3.96
	R	1.20	2.76	2.37	2.15	2.33	2.49	3.23	2.68	3.07
EC	L	2.00	2.06	2.10	3.01	2.33	1.91	2.48	2.45	2.49
	R	1.80	2.42	1.83	3.75	1.64	3.10	2.88	3.87	3.25
CM	L	2.01	2.19	2.91	1.76	2.41	1.67	1.36	1.94	2.15
	R	1.70	1.69	2.15	1.51	1.85	1.76	1.80	1.48	2.78
MH	L	1.19	1.55	2.54	2.01	2.32	2.11	2.76	1.95	2.43
	R	1.14	1.19	1.37	2.06	1.63	1.12	1.33	.90	2.53
RC	L	2.41	2.37	2.46	2.10	3.19	1.94	3.46	2.16	2.33
	R	2.98	2.42	2.97	2.12	1.69	2.68	1.83	2.68	2.28
RG	L	1.18	2.27	1.54	1.14	1.27	1.70	1.70	2.00	3.08
	R	1.68	2.18	.67	.98	1.43	2.61	1.96	2.68	.98
Mid-	L	1.99	2.06	2.04	1.96	2.32	1.94	2.48	2.00	2.49
Score:	R	1.27	1.69	1.44	2.06	1.69	2.36	1.96	2.68	2.53
	L+R	1.39	1.84	1.64	2.00	1.90	2.07	2.40	2.20	2.51

the group within 1.58 — 8.37 dB, the mid-value being 4.28 dB (a tendency exists for the higher frequencies to yield more varied distributions, as would of course be expected.) The group mean function, however, can be stated with a precision ($\pm$ 1 Standard Error) of 0.46 — 2.42 dB, mid-value being 1.14 dB (highs again slightly less precise).

It is impossible from these data to determine the part played in these differences among subjects by (a) real differences in the acoustic coupling of earphones to ears — for some subjects the 705A earphone might be more efficient than the circumaural, for some subjects the reverse, and (b) the totality of variance sources (1) — (8). However, we have estimates of (1) — (4), the effect of earphone-eardrum coupling for the standard earphone. Harris (1954) showed that the SD of ten threshold crossings at all octaves 256-8192 Hz deteriorated by no more than about 1 dB when a comparison was made between the condition of removing-replacing, or not, the earphone after each threshold cross-

TABLE II

DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS DISCRIMINATION

Entry: Standard Deviation in DB of Ten Consecutive Loudness Equality Judgments at 40 DB Sensation Level on Standard Earphone. Comparison Earphone: TRACOR Corp. "Otocup"

Subj.	Stand. Phone On	Frequency in KHz 0.25	0.50	0.75	1	2	3	4	6	8
AR	L	1.62	1.92	2.28	1.91	2.19	1.18	2.32	.70	1.79
	R	1.60	1.25	1.78	1.78	1.20	2.68	4.96	4.17	2.96
JR	R	1.60	1.25	1.78	1.78	1.20	2.68	4.96	4.17	2.96
	R	1.43	1.37	1.44	1.17	1.36	.98	1.33	.81	2.42
JD	L	2.65	1.20	1.70	2.29	1.99	2.68	2.15	2.91	2.02
	R	1.37	1.58	1.14	1.79	1.62	1.41	2.76	2.42	3.67
JW	L	.90	1.00	2.26	2.00	1.90	.94	.92	2.21	2.51
	R	2.66	.81	1.20	.46	.98	1.42	1.83	1.25	2.60
HM	L	1.55	1.43	1.43	2.97	2.49	3.32	3.16	2.05	2.29
	R	2.06	1.73	2.34	1.14	2.15	1.99	3.46	2.77	2.83
CMc	L	1.04	1.22	2.19	.94	1.28	1.43	2.16	1.60	3.00
	R	1.68	.98	1.19	1.60	1.72	2.09	1.96	1.97	2.06
MD	L	1.33	1.37	1.04	2.49	1.72	2.53	2.00	2.43	2.61
	R	1.20	2.00	1.56	1.81	2.19	2.28	2.68	2.76	2.57
DW	L	1.08	1.86	2.42	1.84	2.80	3.61	2.11	2.24	6.26
	R	2.33	2.14	1.66	2.83	2.96	2.15	3.83	3.26	4.49
EC	L	1.11	2.15	1.60	2.66	1.43	4.32	2.16	2.24	1.92
	R	1.47	1.95	2.30	1.74	1.10	1.83	1.02	1.33	2.24
CM	L	2.29	1.47	1.78	1.87	1.30	1.11	2.02	1.42	1.55
	R	1.33	1.64	1.56	1.68	1.49	1.19	1.86	1.20	2.00
MH	L	2.18	2.01	2.49	3.10	1.64	2.64	2.16	3.37	2.73
	R	2.99	2.77	2.40	1.80	1.50	3.27	2.90	2.29	3.47
RC	L	3.74	3.01	2.99	3.03	4.71	5.21	4.46	5.04	6.14
	R	3.78	4.85	3.74	4.67	4.44	4.27	3.80	3.52	3.01
RG	L	2.00	3.23	1.49	3.88	3.10	3.72	4.45	3.74	5.99
	R	1.80	1.86	2.84	1.43	2.53	4.92	2.38	2.43	.83
Mid-	L	1.55	1.47	1.78	2.29	1.90	2.53	2.15	2.24	2.51
Score:	R	1.68	1.73	1.66	1.74	1.62	2.09	2.68	2.42	2.60
	L+R	1.61	1.68	1.74	1.82	1.72	2.27	2.16	2.35	2.55

ing. Hickling (1966) found deterioration of SDs under these conditions of 0.5, 0, 1.35, and 1.32 dB at 1, 2, 6, and 8 kHz respectively. A conservative estimate of the contribution of each of the factors (1) — (4) to unreliability of the transfer function would be about 1 dB with experienced subjects. Thus, in the worst case when all four sources (1, 4) tended in the same direction, two thresholds might seem to differ by 4 dB. Of course, with some types of earphone the fit to the head might be more critical, or more difficult to standardize, and the variance might be larger.

The variance associated with (5) — (6), threshold testing, including in most cases also (1) — (4), has often been assessed (for a review of ten such sets of data see Hickling). The latter gave test-retest audiometry to 60 adults and computed intra-subject SDs of 2.27, 2.23, 3.51, and 3.40 dB at 1, 2, 6, and 8 kHz respectively. These figures are representative of the individual test-retest differences one may expect in establishing the

TABLE III

INDIVIDUAL DIFFERENCES IN THE EARPHONE TRANSFER FUNCTION, AND THE PRECISION OF THE MEAN TRANSFER FUNCTION

Entries: Standard Deviations of Distributions of Individual Transfer Functions, and Standard Errors of the Mean Functions

	Frequency in KHz								
	0.25	0.50	0.75	1	2	3	4	6	8
				"Otocup"					
SD	2.23	1.58	1.98	2.36	3.69	3.54	4.14	5.66	8.37
$S.E._{Mn}$	0.64	0.46	0.57	0.68	1.07	1.02	1.20	1.64	2.42
				"Auraldome"					
SD	5.42	4.43	3.43	3.61	3.75	5.40	4.79	6.82	6.28
$S.E._{Mn}$	1.57	1.28	.99	1.04	1.08	1.56	1.38	1.97	1.82

TABLE IV

DATA ON DISTRIBUTION OF TEST-RETEST DIFFERENCES IN EARPHONE TRANSFER FUNCTIONS BY THE TRADITIONAL LOUDNESS-BALANCING PROCEDURE

Entries: Mean Differences for Individual Transfer Functions, and the Standard Deviation and Standard Error of each Mean Difference

	Frequency in KHz								
	0.25	0.50	0.75	1	2	3	4	6	8
Mean	6.30	6.30	6.46	5.76	5.22	4.16	4.38	7.54	7.16
SD	5.30	4.82	4.24	4.50	4.98	3.74	2.82	5.64	6.12
$S.E._{Mn}$	1.47	1.34	1.18	1.47	1.38	1.04	0.78	1.56	1.70

thresholds of (5) and (6) with experienced subjects such as would be used in deriving earphone transfer functions.

The size, then, of individual differences in transfer functions in Table III, often ± 5 dB or more distant from the mean function, can easily be explained in terms of the variances associated with four couplings, two threshold collections, and two alternate interaural loudness balancings. It is not necessary to invoke much real individual difference in the efficiency of the fit of the earphone to the individual head. This last is of considerable importance when one wishes to explore some procedure which reduces the number of subjects.

C. Individual Test-Retest Differences in Transfer Functions.

The reader has of course already seen that variances (1) — (8) are each in terms of a ± sign, and would tend to cancel each other for the most part. Only an actual replication of the total process will give a realistic estimate of the total variance in the transfer function. In the present data, an estimate of test-retest reliability of the whole transfer procedure was to be had. Test-retest audiometric threshold differences within either cushion are no greater than those between cushion used here when corrected for differences in the sensitivity of the two drivers. Differences were therefore computed between the individual transfer functions from one circumaural earphone to another. Data on these differences are in Table IV. It is seen that test-retest differences between cushion transfer functions are as large as ± 6 dB at about half of the frequencies at random. With an N of 13, this means that a Standard Error of about 1.7 dB is achieved for the mean transfer function. If, however, a Standard Error of no greater than 1.0 dB were desired, as should not be rare, it would have to be recommended that the N be raised to about 35 persons, which would often prove an eco-

nomic deterrent, or else ways sought strenuously for new methods to reduce the number or extent of the eight major sources of variance inherent in this traditional procedure.

EXPERIMENT II: Monaural Loudness Discrimination

It occurred to us that if somehow one could reduce the problem to one of loudness discrimination in one ear, rather than of the somewhat more variable interaural loudness-equality judgment, and avoid the necessity altogether of taking absolute threshold judgments, with their variance each of several dB, a distinct gain in reliability for the same expenditure of time and energy could be expected. A hint was provided by the technique used by the Physikalische-Technische Bundesanstalt in West Germany (Weissler, 1968); an earphone is placed on one ear throughout, and another standard and the unknown phone are placed in succession on the other ear. Thus, one avoids the matter of differences in acuity between the ears, since the standard and all unknown earphones are applied to the same test ear. However, it still incorporates interaural loudness balancing.

Our solution was to create a constant-level tone in the test ear with a bone-conduction vibrator on the forehead and an appropriate masking noise in the non-test ear. This tone is placed at, say 40 dB sensation level, but its exact SL is irrelevant, so long as for all earphone comparisons it remains constant. This bc tone is then pulsed alternately with an air-conduction tone from the standard earphone and subsequently with any other phone of interest. The bc stimulus thus serves as a constant reference loudness against which the outputs of all phones can be compared. The voltage of a new earphone is simply compared with that from the standard, and the differences used to write new standard voltages for the new phone over the audiometric frequency range. The variances in this procedure are simply those associated with coupling the standard and unknown earphones to the eardrum, and that associated with two monaural loudness discriminations, for a total of four sources.

The mean differential sensitivity (JND) by the Method of Adjustments for monaural loudness discrimination has been estimated as 1.23 — 1.61 dB for seven subjects over the range 125 — 6000 Hz at 1 sone. (Harris, 1963).

For exploring this monaural loudness-

TABLE V

TEST-RETEST DIFFERENCES IN INDIVIDUAL TRANSFER FUNCTIONS

Comparison Earphone: TRACOR "Otocup"

Frequency in KHz

	0.25			0.5			1			2			4			6		
Subj	**T**	**Re-T**	**D**	**T**	**Re-T**	**D**	**T**	**Re-T**	**D**	**T**	**Re-T**	**D**	**T**	**Re-T**	**D**	**T**	**Re-T**	**D**
MD	3.0	6	3	10.5	10.5	0	7.5	9	1.5	4	10.5	6.5	20.5	16.5	4	15	8	7
JD	0.5	1.5	1	8.5	5.5	3	11	14	3	6	1	5	6	15	9	3.5	0.5	3
JH	3	2.5	0.5	7.5	7	0.5	12	11.5	0.5	5	8	3	15.5	8.5	7	19	15	4
CMc	—4	5	9	8.5	2.5	6	6	5	1	12	8.5	3.5	11	6	5	16.5	12	4.5
VM	2.5	—1	3.5	8.5	9	0.5	7	8.5	1.5	6.5	7	0.5	23	17.5	5.5	22.5	26.5	4
CM	7.5	7.5	0	1	5.5	4.5	4.5	8	3.5	6.5	7	0.5	17.5	17.5	0	21.5	25.5	4
JR	—6	—10.5	4.5	2	2	0	9.5	10.5	1	10	6	4	17	13.5	3.5	12.5	10.5	2
JS	—6.5	—7	0.5	19	24	5	0.5	2	1.5	9	18	9	12.5	14.5	2	14.5	20	5.5
FW	—9	—1.5	7.5	13.5	15	1.5	8	4.5	3.5	16	11	5	12.5	18.5	6	26.5	34	7.5
Mn T	—1			8.8			7.3			8.3			15.0			16.8		
Mn ReT		0.3			9.0			8.1			8.5			14.2			16.9	
Mean Diff:			3.28			1.33			1.89			4.11			4.67			4.61
$S.E._{Mn\ Diff}$			1.07			0.86			0.35			0.90			0.89			0.59
T_{Mn}—Re-T_{Mn}		1.3			0.2			0.8			0.2			0.8			0.1	

TABLE VI
TEST-RETEST DIFFERENCES IN INDIVIDUAL TRANSFER FUNCTIONS
Comparison Earphone: Maico "Auraldome"
Frequency in KHz

	0.25			0.5			1			2			4			6		
Subj	T	Re-T	D	T	Re-T	D	T	Re-T	D	T	Re-T	D	T	Re-T	D	T	Re-T	D
MD	5	12	7	13	8	5	10	4	6	6.5	9.5	3	14	14.5	0.5	18	12	6
JD	9	1	8	6	3	3	8.5	4.5	4	5	2	3	12	12	0	10	2	8
JH	1.5	6.5	5	1	4	3	7	10	3	5.5	7.5	2	10	3	7	21.5	15	6.5
CMc	8	2.5	5.5	9	8.5	0.5	3.5	2.5	1	5.5	1	4.5	13.5	7.5	6	10	12.5	2.5
VM	1	1	0	11.5	7.5	4	2	1.5	0.5	5.5	7.5	2	23.5	20.5	3	17.5	13	4.5
CM	10	5	5	2.5	2	0.5	5.5	0.5	5	9	11	2	17.5	19	1.5	28	21.5	6.5
JR	—1	—6.5	5.5	8	5	3	—0.5	5	5.5	10	8	2	19.5	14	5.5	16	9.5	6.5
JS	2	0.5	1.5	8.5	11.5	3	—5.5	—8	2.5	9	9.5	0.5	13.5	14	0.5	21	22	1
FW	—1.5	—2	0.5	14	14	0	7.5	4	3.5	0.5	0.5	0	21	24.5	3.5	24	35.5	11.5
Mn T	3.8			6.4			4.2			6.3			16.05			18.4		
Mn Re-T		2.2			5.9			2.7			6.3			14.3			15.9	
Mn Diff:			4.22			2.44			3.44			2.33			3.06			5.89
$S.E._{Mn\ Diff}$			0.95			0.57			0.64			0.45			0.88			1.01
T_{Mn}—Re-T_{Mn}		1.6			0.5			1.5			0			1.75			2.5	

discrimination procedure for calibrating earphones, eight graduate students in sensory psychophysiology were used, all with normal hearing, and two older psychoacousticians with some mild high-frequency hypacusis. Individuals were seated in a double-walled audiometric chamber of 600 sq. ft. lined with 4-inch fiberglass batts. All equipment except earphones and hand switch was in an adjoining control space.

The output of a General Radio Type 1304 pure-tone generator was split and led to (I) a bc vibrator, and (II) an earphone. Channel I was led to one channel of a Grason-Stadler Model 829S71 electronic switch, a 1-dB/step attenuator, a Hewlett-Packard Model 465A amplifier, and finally to a Radioear Model B70A bc vibrator. The vibrator was fuzed to a 1-inch wide flexible band stretched firmly around the head of the subject, the vibrator resting on the middle of the forehead.

Channel II was led to a rotary attenuator and paper-tape voltage recorder constructed on the Bekesy-tracking principle, through a second Grason-Stadler Model 829S71 switch, and to any of three ear phones.

The two switches were driven by a pair of Grason-Stadler Model 471 timers connected so that Channels I and II could be alternated with any desired timing. All rise-fall times were 40 msec. The bc tone was on for 0.4 sec, the ac for 0.6 sec. Intervals between the two were at first set at 40 msec; with this pattern the subject experienced a shorter tone alternating with no appreciable pause with a longer tone, both in the same ear. The effect was thus of monaural loudness discrimination: at equal loudness, the subject heard an almost uninterrupted pure tone of constant loudness, and this last judgment could be made with great surety.

However, with the constantly-changing intensity of the ac channel inherent in the Bekesy tracking, this loudness equality is always being upset, and subjects not rarely lost track of whether the ac or the bc tone was weaker, and uncertainty existed as to whether the ac signal, which the subject controlled with a microswitch, should be made louder or softer. In order to correct this, the interval between tones was increased to 300 msec, and the interval between bc-ac pairs to 1 sec. With this pattern, subjects were never confused as to which direction the ac tone should be changed, and at equal loudness the experience was of a monaural train of pairs of pure tones of somewhat unequal length, but all of the same quality and loudness.

A Western Electric 705A earphone served as standard, against which were loudness-balanced a Maico Co. "Auraldome" and a TRACOR Corp. "Otocup" circumaural earphones, each fitted with a Permoflux Corp. PDR-600 driver. Each phone was in an appropriate commercial headband; on the other side was a suitable earphone delivering a third-octave band of noise from a Beltone masking generator set to an effective masking level of at least 40 dB.

The experimenter seated the subject, fitted the headband, and adjusted one of the three earphones on the test ear. An appropriate masking noise was applied to the non-test ear, whereupon a bc threshold was taken by the Method of Limits at one of the frequencies. This bc sound was of course referred to the nonmasked test ear. The bc stimulus was increased by 40 dB and subject asked to increase the ac signal, using his Bekesy-type hand microswitch, to yield equal loudness between the bc and ac signals, and thereafter to track loudness equality for a couple of minutes.

Frequencies were introduced in random order within subjects, and earphones in random order across subjects. Finally, the same bc reference intensity would create different loudnesses depending upon the occlusion effect of large or small earphone/cushion cavities, at the lower frequencies a tight-fitting wax-impregnated earplug was sealed into the test ear meatus to eliminate the occlusion effect from perturbing the data by maximizing it across all phones.

With the situation maximized by using earplugs where called for, and increased intervals between tones where subjects requested it, Tables V-VI show the individual test-retest differences between (1) an initial standard-unknown earphone comparison, and (2) the same comparison resulting from a later complete replication of the whole set of judgments. It is from the distributions of these individual differences that we can assess the general reliability of the procedure.

The tables show that the average subject yields a test-retest difference of from 1.33—5.89 dB, mid-value of 3.34 dB. As usual, the lowest and highest frequencies show the larger differences. These mean test-retest differences can be compared with those of Willott and Myers in Table IV above for the identical earphones; they are about half as large as with the traditional procedure, with Standard Error proportionately small. The consistency of the individual in test-retest would seem adequate for most purposes, and reflects largely the variance associated with fitting the earphones to the head. The size of the sample here would seem a minimum for assessing that variance.

The reliability of the group means is shown by a comparison of mean test-retest voltages. These differences are included in the last row of each table. On a complete replication, this group duplicated its mean transfer function within a dB or two (mid-value of test-retest differences = 0.83, range of 0—2.5 dB).

The final means for each frequency, when corrected for the efficiency of the drivers, can be used to write a new standard reference equivalent threshold SPL for each of the two earphones used, but we are not here concerned with that detail.

Some of the efficiencies inherent in this new procedure are apparent, such as the usability of most available subjects, not just those with normal and nearly symmetrical hearing, or at least free from recruitment, the ease with which subjects can make loudness discriminations rather than loudness balances, and the time consumed. We have regularly found that all observations on a subject at a particular frequency can be completed in five minutes; in the traditional procedure 20 minutes would be a minimum.

We may conclude that acceptable group means for transferring audiometric standards from a known to an unknown earphone would be obtained at any frequency by requiring as few as nine subjects to make a single monaural loudness discrimination per phone by this technique.

REFERENCES

1. Anon, ISO Recommendation R389, "Standard Reference Zero for the Calibration of Pure-Tone Audiometers." International Standards Organization. Nov. 1964 (Available through the USA Standards Institute.)

2. Pearl G. Weissler. "International Standard Reference Zero for Audiometers." J. Acoust. Soc. Amer., 44, 264-275 (1968).
3. J. A. Willott, C. K. Myers, and J. D. Harris, "Differential Sensitivity for Alternate Interaural Loudness Balancing in the Psychoacoustic Calibration of Earphones." USN Submarine Medical Center Report, 1969 (In Press).
4. Virginia Morse and J. D. Harris, "A Monaural Comparison of Two Circumaural Earphones with a Standard Audiometric Earphone." USN Submarine Medical Center Report, 1969 (In Press).
5. J. D. Harris, "Loudness Discrimination." J. Speech Hear. Dis., Suppl. 11 (1963).
6. R. C. Rowland, and J. V. Tobias, "Interaural Difference Limen." J. Speech Hear. Res., 10, 745-756 (1967).
7. J. F. Jerger and E. R. Harford, "Alternate and Simultaneous Binaural Balancing of Pure Tones." J. Speech Hear. Res., 3, 15-30 (1960).
8. J. D. Harris, "Experiments on Fluctuations of Auditory Acuity." J. Gener. Psychol., 50, 87-109 (1954).
9. S. Hickling, "Studies on the Reliability of Auditory Threshold Values." J. Audit. Res., 6, 39-46 (1966).

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* NAVAL SUBMARINE MEDICAL CENTER, Submarine Medical Research Laboratory	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

AN EFFICIENT MONAURAL PROCEDURE FOR THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Interim Report

5. AUTHOR(S) *(First name, middle initial, last name)*

J. Donald HARRIS

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
5 September 1969	9	9

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. MF12.524.004-9010D.05	SMRL Report Number 593
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Submarine Medical Center Box 600 Naval Submarine Base Groton, Connecticut 06340

13. ABSTRACT

A new procedure is examined for psychoacoustic calibration of earphones in which the air-conducted outputs of a standard and an unknown earphone are successively equated for loudness to a reference bone-conducted tone. The problem to the subject is one of monaural loudness discrimination, with a relatively small variance (differential sensitivity = 1.23 - 1.61 dB), and involves only four sources of variance associated with coupling two earphones to the same ear, and a single loudness discrimination judgment for each phone. The mean test-retest difference in the earphone transfer functions varied by 1.33 - 5.89 dB at different frequencies, mid-value = 3.36 dB. Only a few minutes are required to complete a subject's observations at any frequency. Acceptable group means for transferring audiometric standards to an unknown earphone could be obtained at any frequency by requiring as few as nine subjects to make a single monaural loudness discrimination per earphone by this technique. In contrast, the traditional alternate interaural loudness balancing has a somewhat larger variance associated with the judgment *per se* (differential sensitivity = 1.50 - 2.50 dB), and involves the additional variances associated with collecting two absolute thresholds and coupling the two earphones to a second ear when the earphones are reversed on the head. The mean test-retest difference in the earphone transfer functions by the traditional "ear-reversal" method varied from 4.16 - 7.54 dB at different frequencies, mid-value = 6.30 dB, nearly twice that of the suggested procedure.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Earphones, calibration of (new method)						
Loudness discrimination						
Psychoacoustic calibration (for earphones)						